

[Open in app](#)

# Parveen Khurana

124 Followers

[About](#)[Following](#)

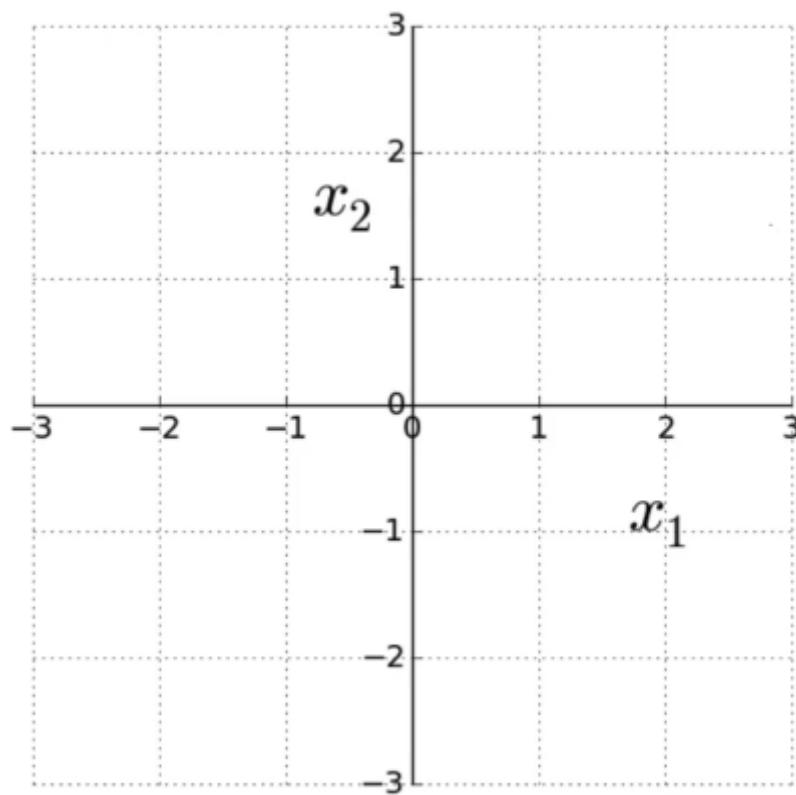
## Vectors and Matrices



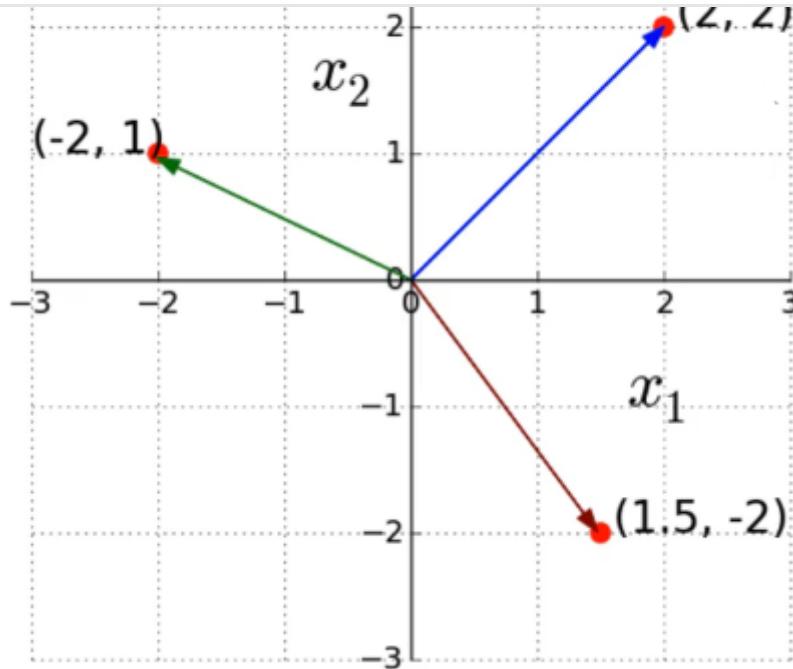
Parveen Khurana Nov 29, 2019 · 11 min read

In this article, we discuss the basic operations on Vectors and Matrices and how we can represent the input data in the form of Matrices.

Vector just represents the co-ordinates that the point has in a given space.



Vectors also have a geometric interpretation: it's a ray that connects the origin of that space to the point.

[Open in app](#)

A vector is quantified with **magnitude** and **direction**.

- **Magnitude of a vector**
- **Direction of a vector**

### Magnitude of a vector

$$\sqrt{(x_1^2 + x_2^2)}$$

We can have vectors which have different directions but the same magnitude

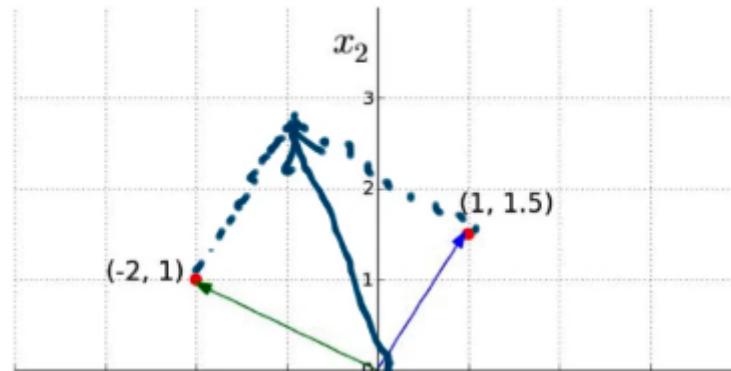
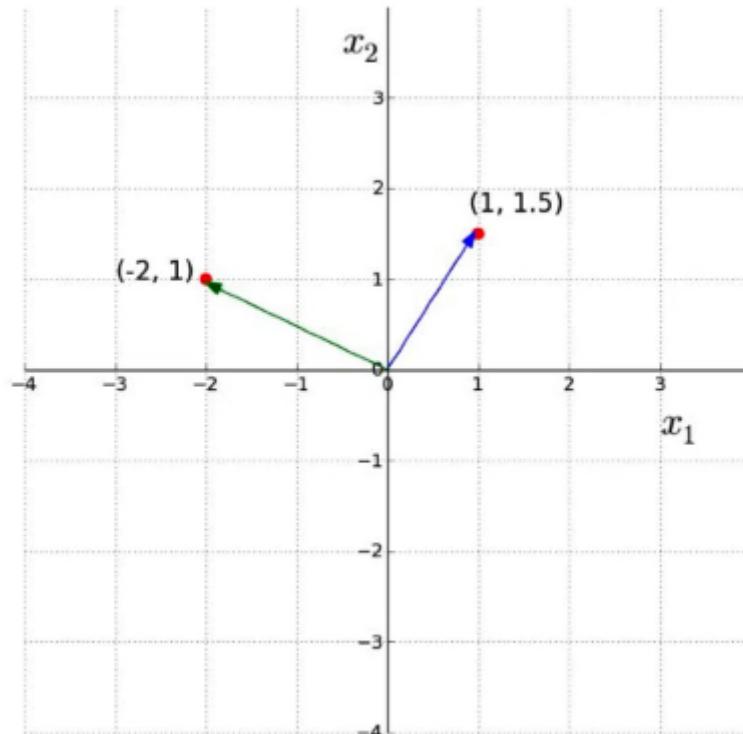
### Euclidean Space

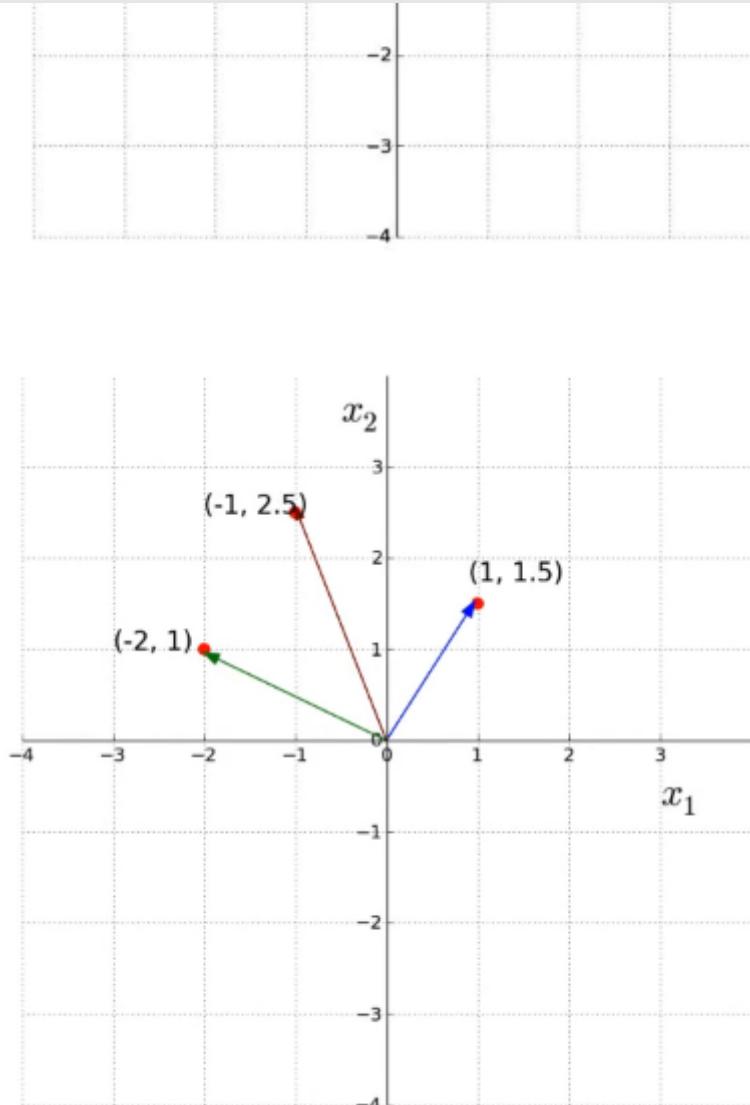
In geometry, a two- or three-dimensional **space** in which the axioms and

[Open in app](#)

is a **Euclidean space**. A **space** in any finite number of dimensions, in which points are designated by coordinates (one for each dimension) and the distance between two points is given by a distance formula. **L2 Norm is also called Euclidean norm.**

**Adding two vectors:** We can add them elementwise

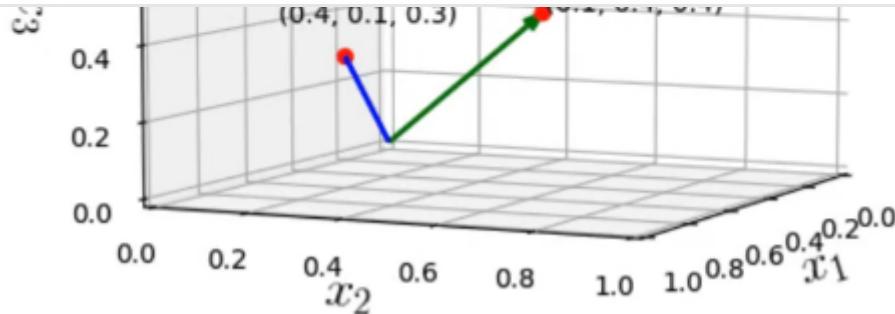


[Open in app](#)

## Addition

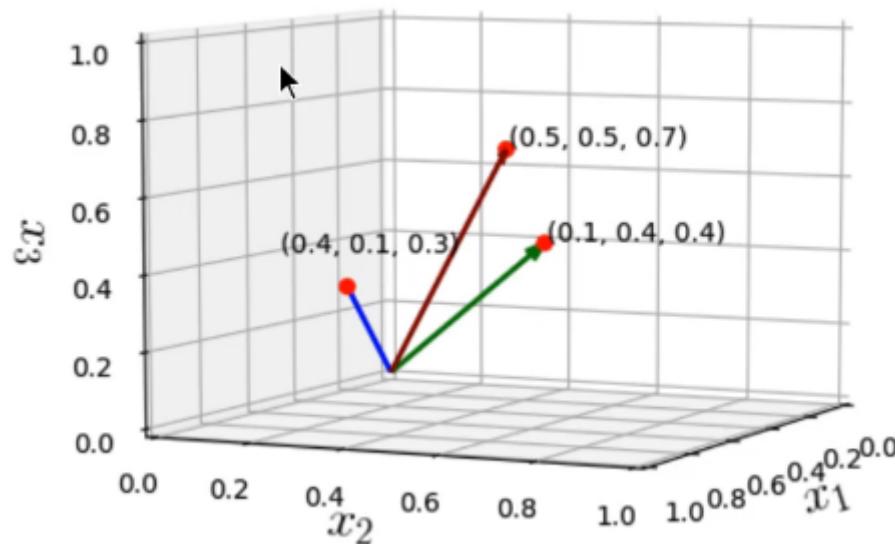
$$\begin{bmatrix} -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -1 \\ 2.5 \end{bmatrix}$$



[Open in app](#)

## Addition

$$\begin{bmatrix} 0.1 \\ 0.4 \\ 0.4 \end{bmatrix} + \begin{bmatrix} 0.4 \\ 0.1 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.7 \end{bmatrix}$$

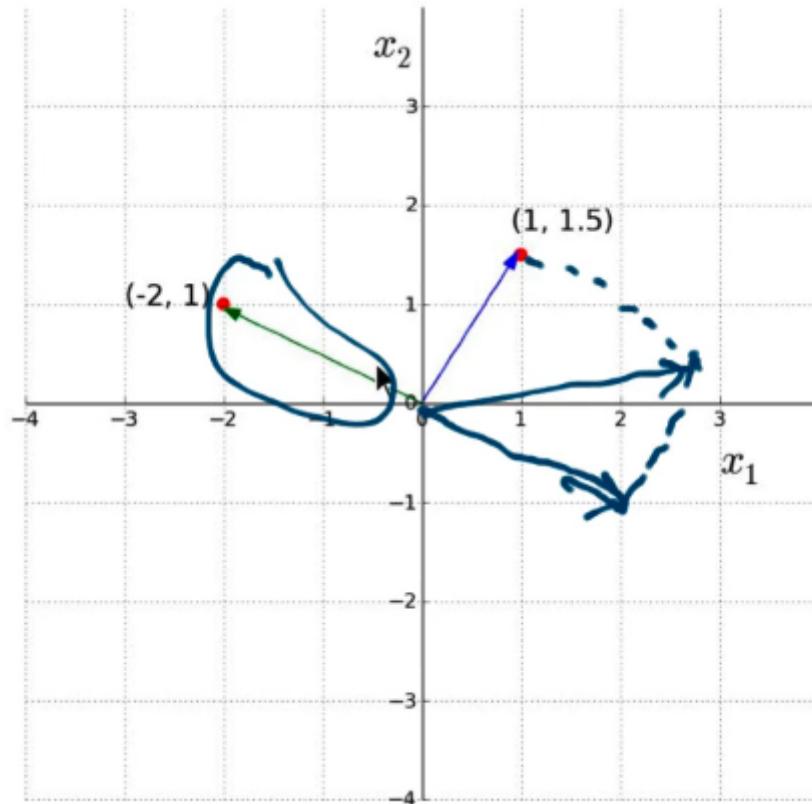


Subtracting two vectors:

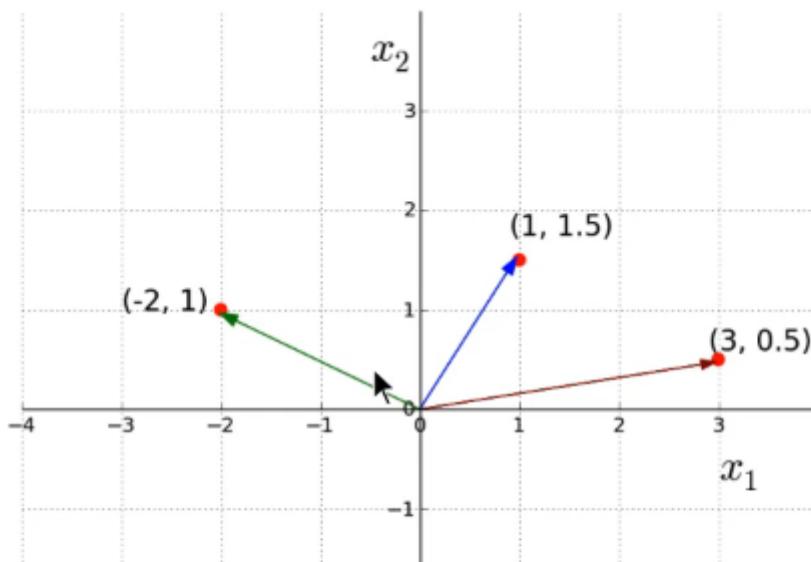
## Subtraction


[Open in app](#)

$$\begin{bmatrix} 1.5 \end{bmatrix} - \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 0.5 \end{bmatrix}$$



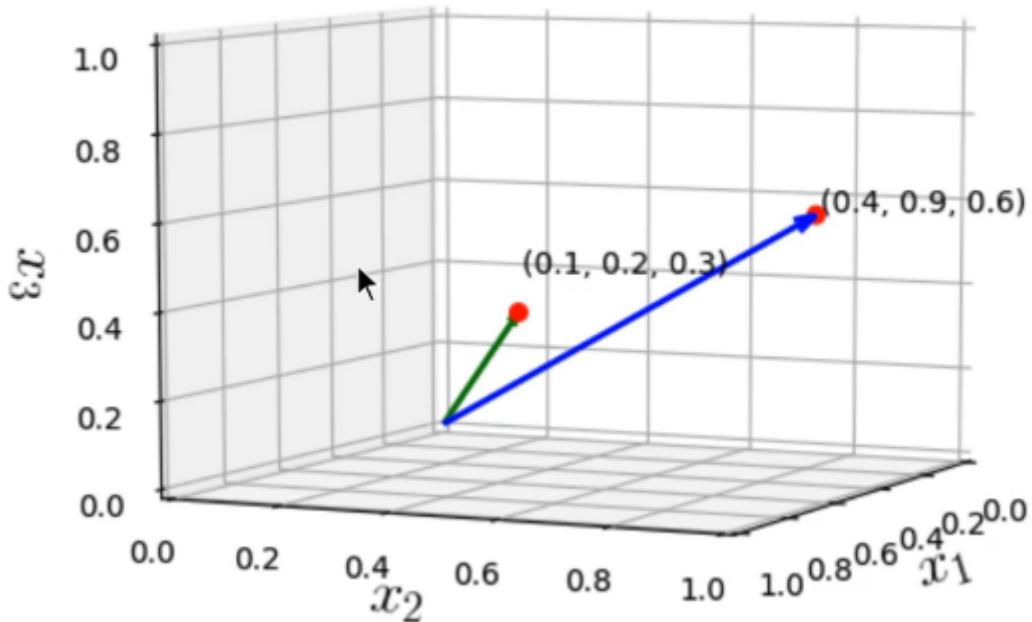
We take the negative of the vector and then it's just simple addition.



[Open in app](#)

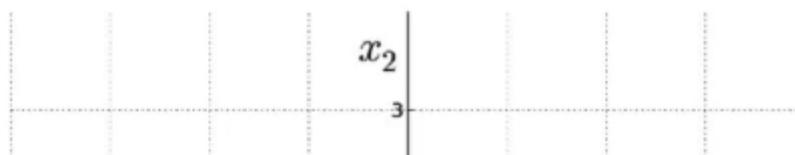
## Subtraction

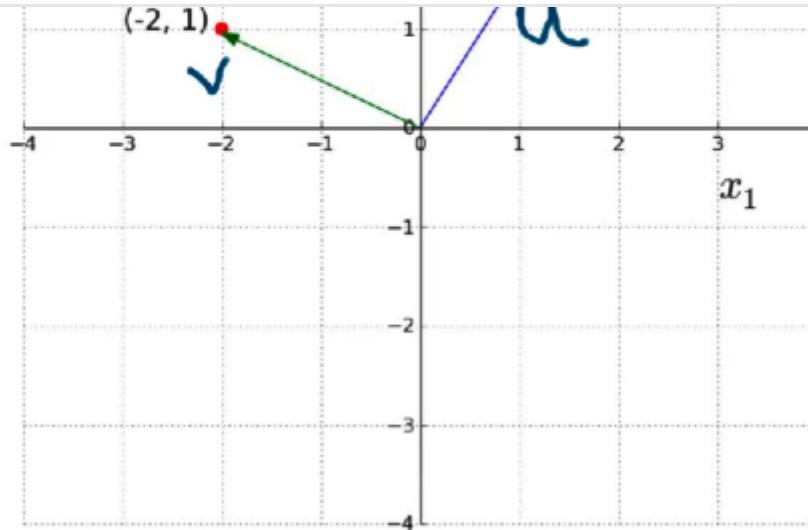
$$\begin{bmatrix} 0.4 \\ 0.9 \\ 0.6 \end{bmatrix} - \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.7 \\ 0.3 \end{bmatrix}$$



## Dot Product of Vectors

Dot Product is the element-wise product of two vectors but the resulting quantity is a scalar.




[Open in app](#)


$$u = [u_1, u_2]$$

$$v = [v_1, v_2]$$

**u1** is the coordinate of **u** vector along the first dimension, **u2** is the coordinate of **u** vector along the second dimension.

Similarly, for **n** dimensions, we write it as:

$$u = [u_1, u_2, \dots, u_n] \in \mathbb{R}^n$$

The two vectors must have the same dimensions (to compute the dot product)


[Open in app](#)

$$= \sum_{i=1}^n u_i N_i$$

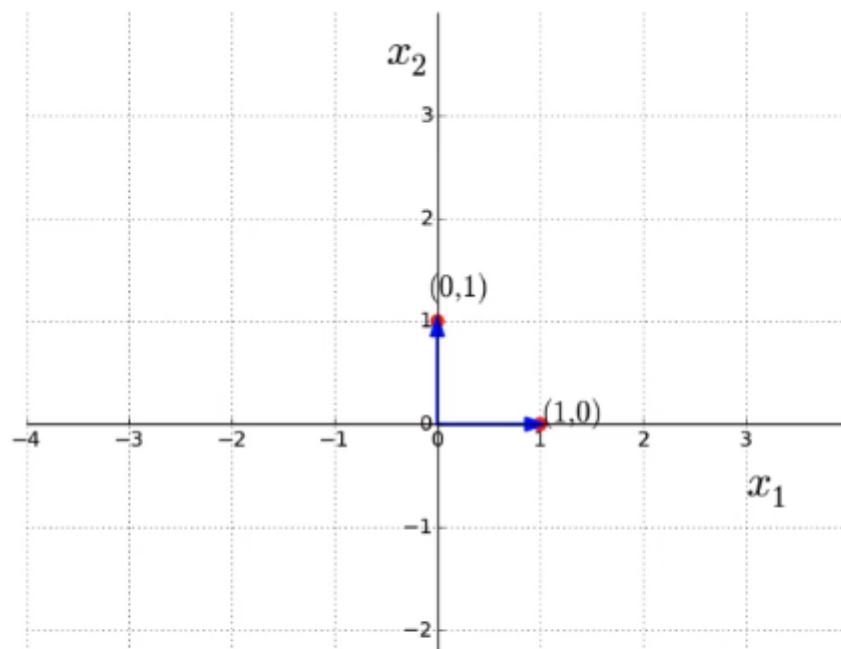
## Multiply

$$\begin{bmatrix} -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} = 0.5$$

The answer, in this case, would be negative of 0.5

## Unit Vector

Any vector of magnitude 1 is called a unit vector.

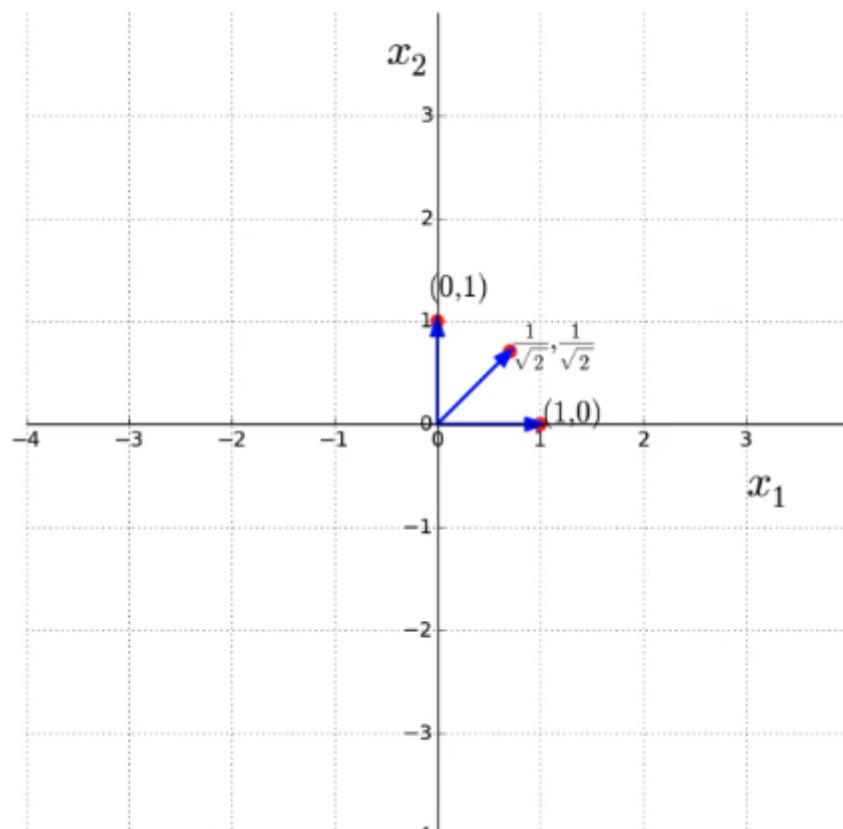


[Open in app](#)

*"Any vector of  
magnitude 1?"*

$$\|(0, 1)\| = \sqrt{(0^2 + 1^2)} = 1$$

$$\|(1, 0)\| = \sqrt{(1^2 + 0^2)} = 1$$

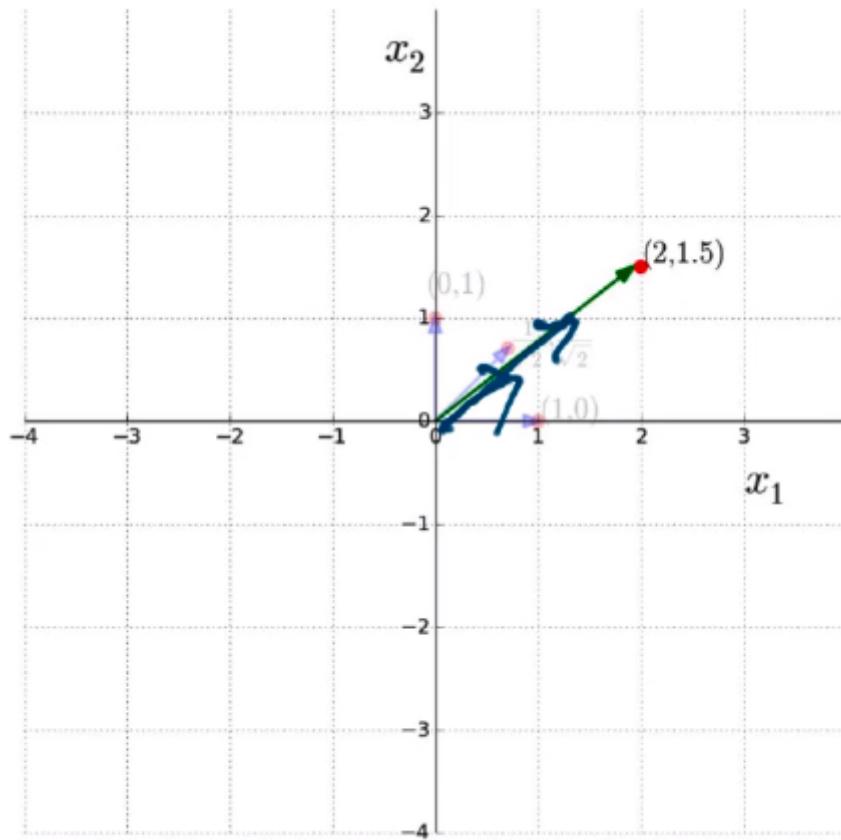


[Open in app](#)

$$\left\| \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \right\| = \sqrt{\left( \frac{1}{\sqrt{2}} \right)^2 + \left( \frac{1}{\sqrt{2}} \right)^2} = 1$$

The above one is also a unit vector, it's not along the direction of the  $x_1$  axis or  $x_2$  axis.

Computing a unit vector in the direction of a given vector:



The above vector  $(2, 1.5)$  is not a unit vector and we want to find a unit vector in its direction.

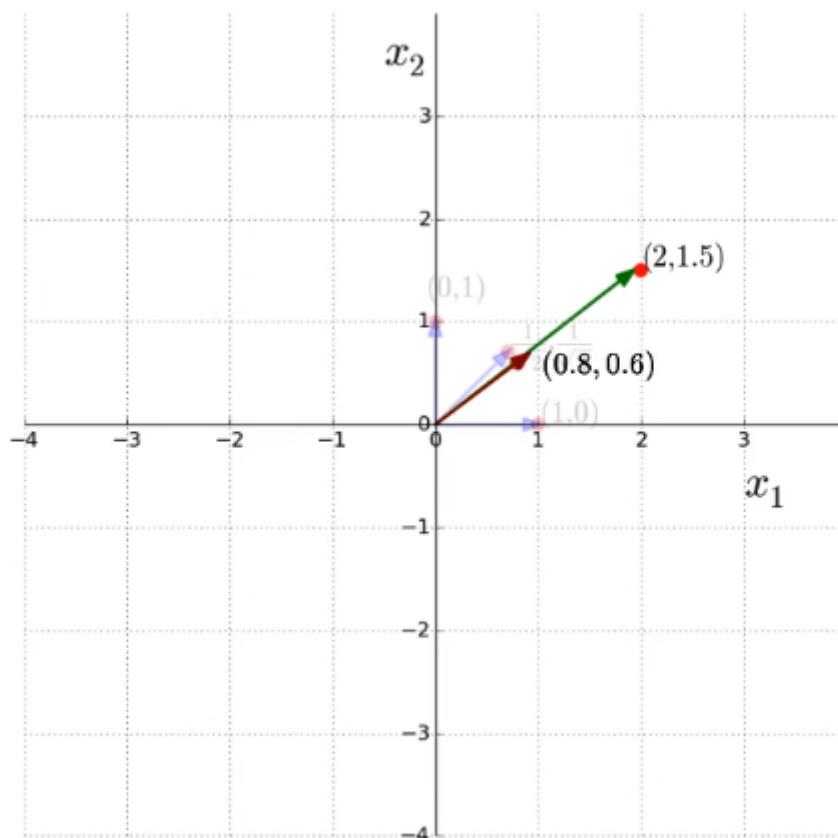
$$\left\| (2, 1.5) \right\| = \sqrt{2^2 + (1.5)^2} = \sqrt{6.25} = 2.5$$

[Open in app](#)

$$\|(2, 1.5)\| = \sqrt{(2^2 + (1.5)^2)} = \sqrt{6.25} = 2.5$$

$$u(2, 1.5) = \frac{(2, 1.5)}{2.5}$$

$$\frac{1}{2.5} \begin{bmatrix} 2 \\ 1.5 \end{bmatrix}$$



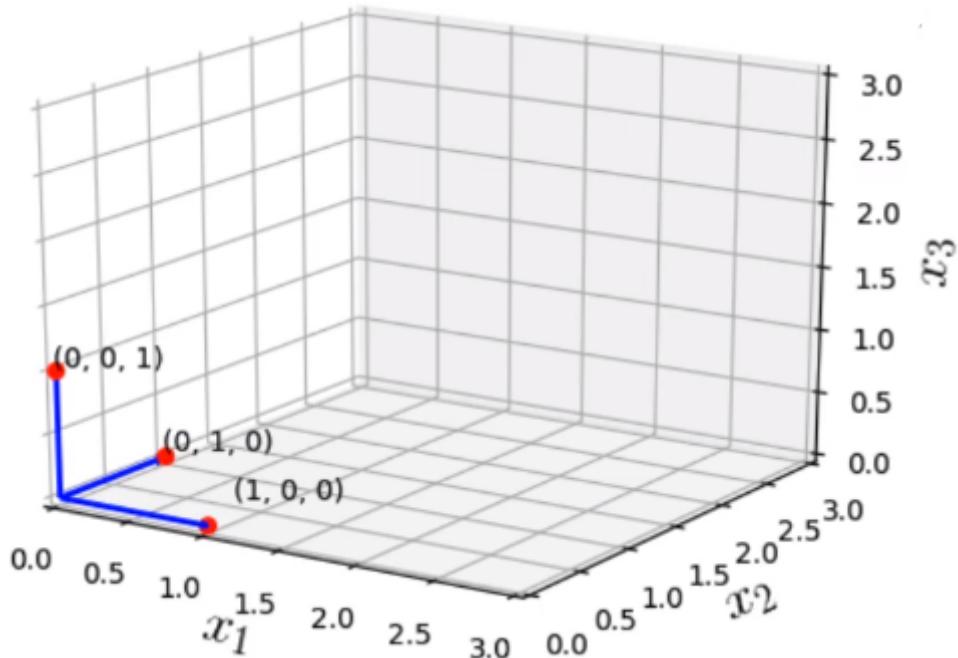
$$\|(2, 1.5)\| = \sqrt{(2^2 + (1.5)^2)} = \sqrt{6.25} = 2.5$$

$\therefore (2, 1.5) \rightarrow (2, 1.5)$

[Open in app](#)

(-1, -1), (2.5, 2.5), (1.5, 1.5)

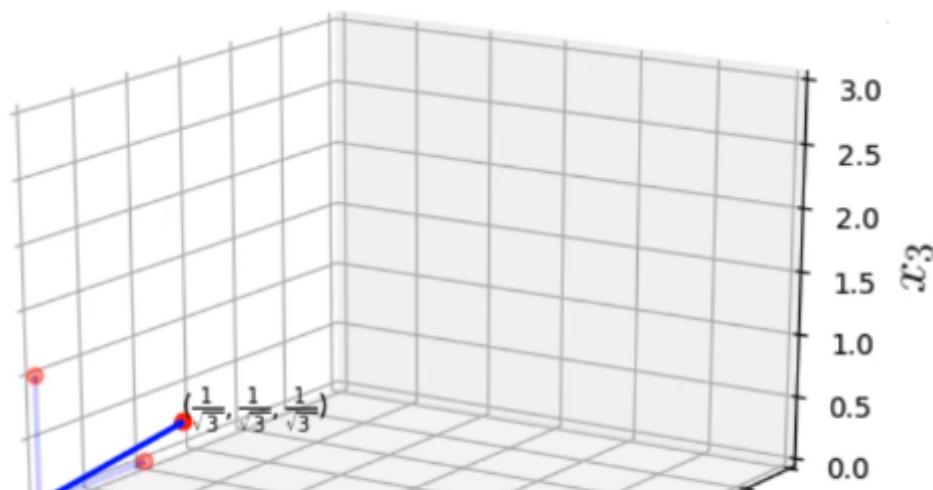
And the same holds in 3 dimensions as well



$$\|(0, 0, 1)\| = \sqrt{0^2 + 0^2 + 1^2} = 1$$

$$\|(1, 0, 0)\| = \sqrt{1^2 + 0^2 + 0^2} = 1$$

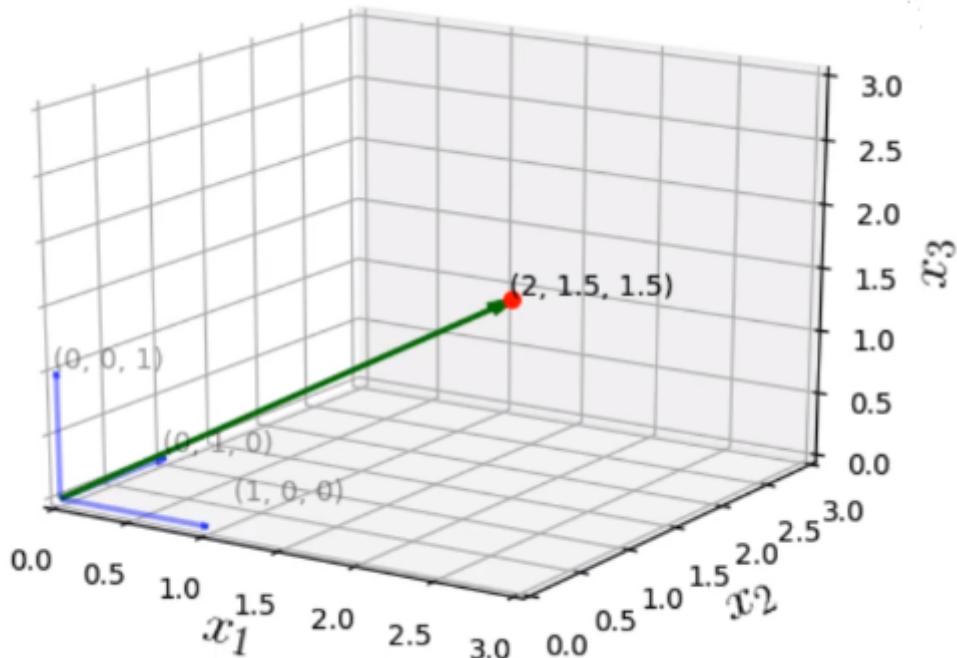
$$\|(0, 1, 0)\| = \sqrt{0^2 + 1^2 + 0^2} = 1$$



[Open in app](#)

$$x_1 \quad -2.0 \quad 2.5 \quad 3.0 \quad 0.0 \quad 0.5$$

$$\left\| \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) \right\| = \sqrt{\left( \frac{1}{\sqrt{3}} \right)^2 + \left( \frac{1}{\sqrt{3}} \right)^2 + \left( \frac{1}{\sqrt{3}} \right)^2} = 1$$



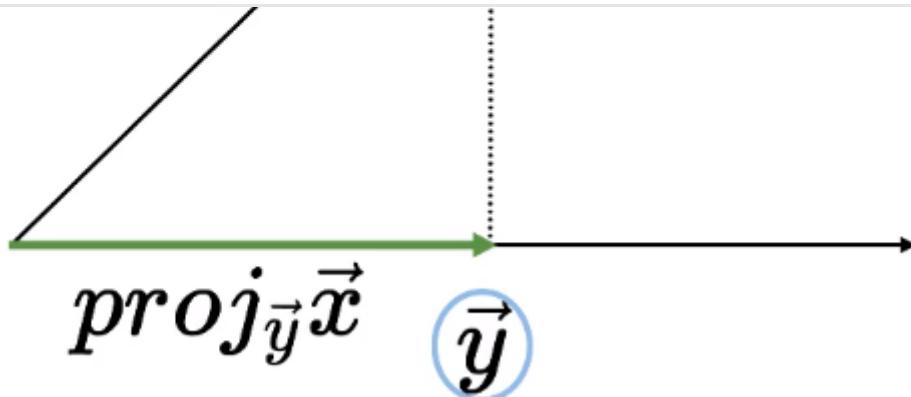
$$\left\| (2, 1.5, 1.5) \right\| = \sqrt{2^2 + (1.5)^2 + (1.5)^2} = \sqrt{8.5}$$

$$u(2, 1.5, 1.5) = \frac{(2, 1.5, 1.5)}{\sqrt{8.5}}$$

$$u(2, 1.5, 1.5) = \left( \frac{2}{\sqrt{8.5}}, \frac{1.5}{\sqrt{8.5}}, \frac{1.5}{\sqrt{8.5}} \right)$$

## Projection of vector onto another vector



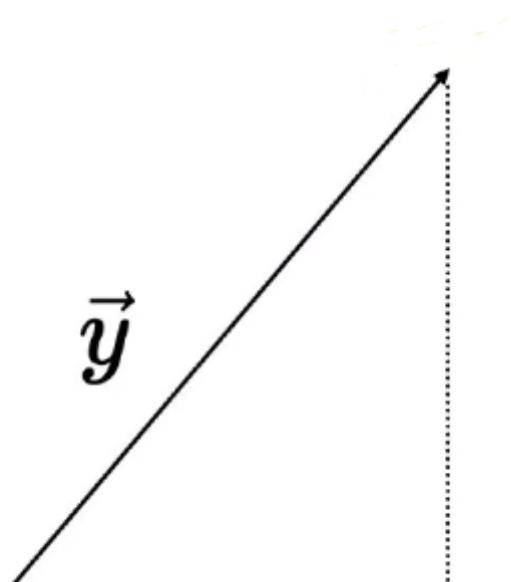
[Open in app](#)

Let's say we have two vectors  $\mathbf{x}$  and  $\mathbf{y}$  and we want to find out the projection of  $\mathbf{x}$  on  $\mathbf{y}$ .

As is clear from the above image, this projection is going to have the same direction as the vector  $\mathbf{y}$ , its magnitude may be different than the vector  $\mathbf{y}$ .

$$\text{proj}_{\vec{y}} \vec{x} = \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right) \vec{y}$$

In this case, the vector was shortened after projection but it is possible that after projection we get back a longer vector than the original one.

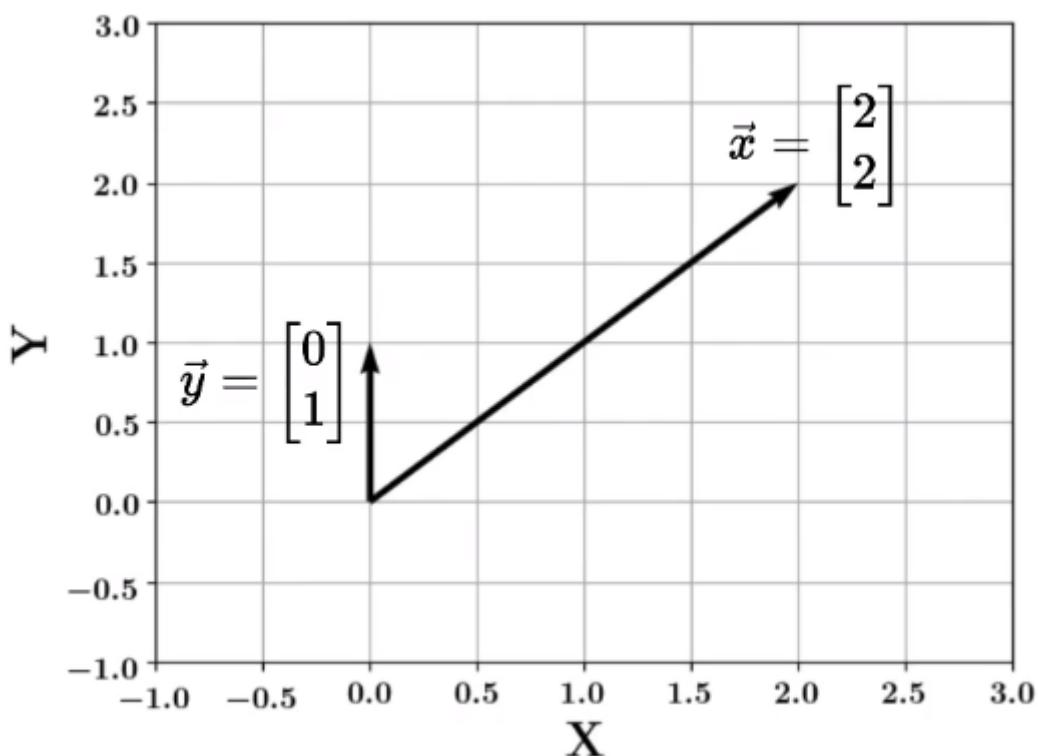


[Open in app](#)

$$\vec{x} \quad proj_{\vec{x}} \vec{y}$$

$$proj_{\vec{x}} \vec{y} = \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|^2} \right) \vec{x}$$

Example:



$$proj_{\vec{y}} \vec{x} = \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right) \vec{y}$$

[Open in app](#)

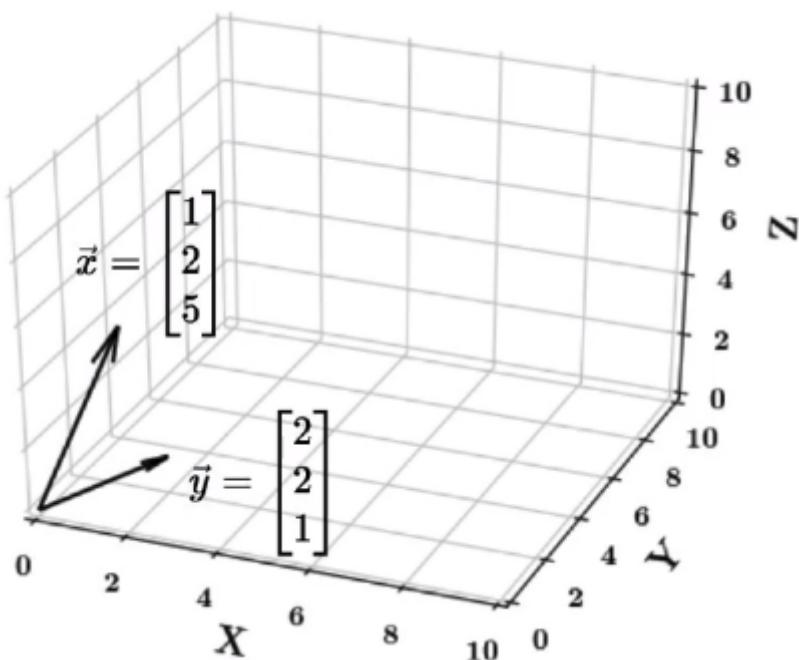
$$\|y\|^2 = \sqrt{0^2 + 1^2} = 1$$

$$\text{proj}_{\vec{y}} \vec{x} = \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right) \vec{y}$$

$$= \begin{pmatrix} \frac{2}{1} \end{pmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

In 3D



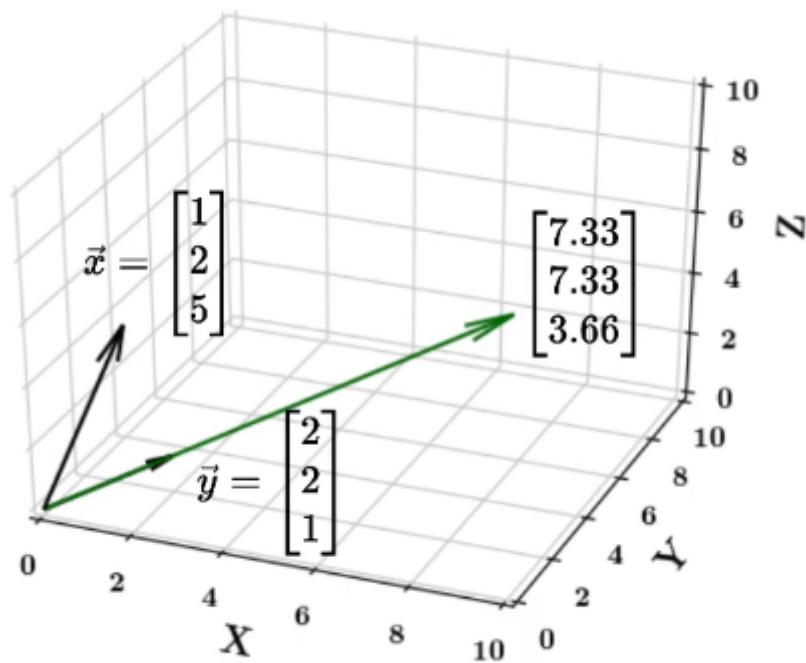
[Open in app](#)

$$\|y\|^2 = \sqrt{2^2 + 2^2 + 1^2} = 3$$

$$\text{proj}_{\vec{y}} \vec{x} = \left( \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2} \right) \vec{y}$$

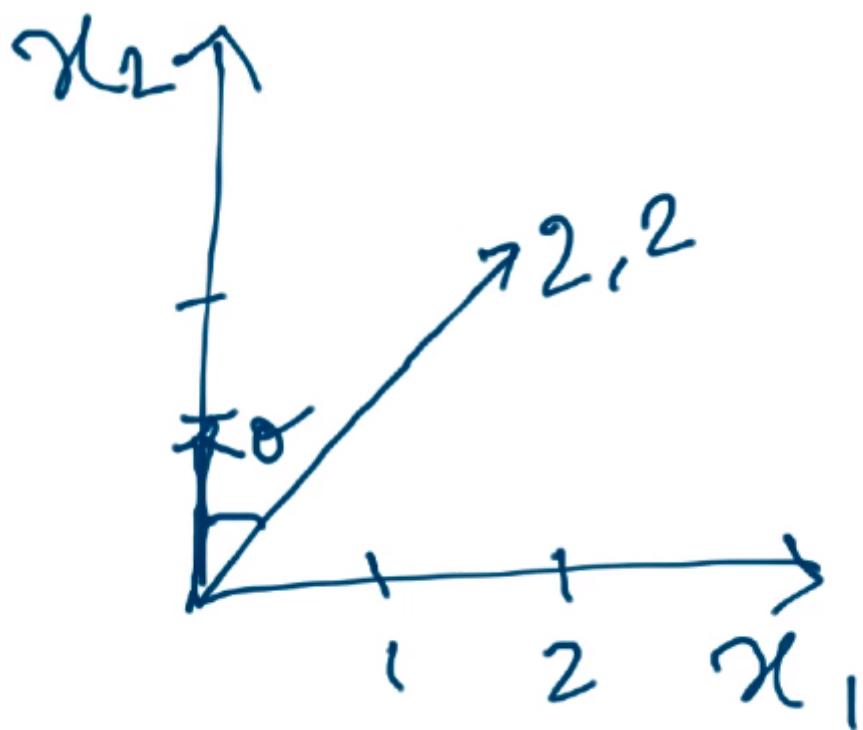
$$= \left( \frac{11}{3} \right) \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 7.33 \\ 7.33 \\ 3.66 \end{bmatrix}$$



[Open in app](#)

$$\vec{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \theta = ?$$



$$\vec{x} \cdot \vec{y} = ||\vec{x}|| ||\vec{y}|| \cos \theta$$
$$\cos \theta = \frac{\vec{x} \cdot \vec{y}}{||\vec{x}|| ||\vec{y}||}$$

$$\vec{x} \cdot \vec{y} = 2 \times 0 + 2 \times 1 = 2$$

[Open in app](#)

$$|\vec{y}| = \sqrt{0^2 + 1^2} = \sqrt{1}$$

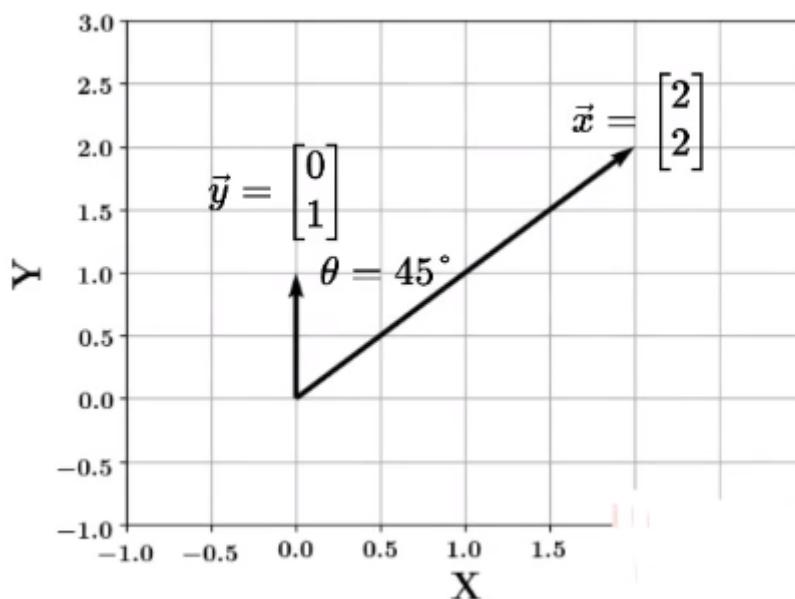
$$\cos\theta = \frac{1}{\sqrt{2}}$$

$$\vec{x} \cdot \vec{y} = 2 \times 0 + 2 \times 1 = 2$$

$$|\vec{x}| = \sqrt{2^2 + 2^2} = \sqrt{8}$$

$$|\vec{y}| = \sqrt{0^2 + 1^2} = \sqrt{1}$$

$$\theta = 45^\circ$$



The same formula can be applied for any number of dimensions.

$\begin{bmatrix} 1 \end{bmatrix}$

$\begin{bmatrix} 2 \end{bmatrix}$

[Open in app](#)

$$\begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

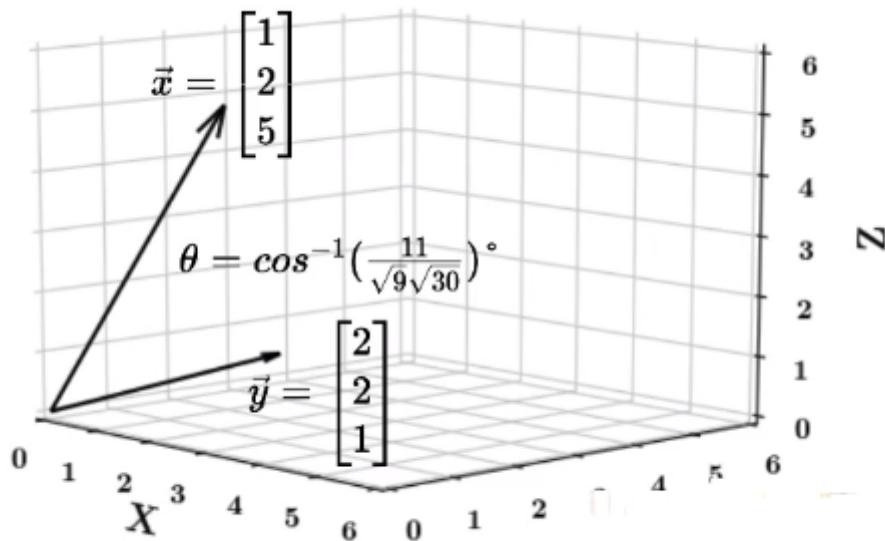
$$\vec{x} \cdot \vec{y} = 1 \times 2 + 2 \times 2 + 5 \times 1 = 11$$

$$|\vec{x}| = \sqrt{1^2 + 2^2 + 5^2} = \sqrt{30}$$

$$|\vec{y}| = \sqrt{2^2 + 2^2 + 1^2} = \sqrt{9}$$

$$\cos\theta = \frac{11}{\sqrt{30}\sqrt{9}}$$

$$\theta = \cos^{-1}\left(\frac{11}{\sqrt{9}\sqrt{30}}\right)^\circ$$



Orthogonal Vectors:

[Open in app](#)

$$\Rightarrow \theta = 90^\circ$$

$$\Rightarrow \cos\theta = \cos(90^\circ)$$

$$\Rightarrow \cos\theta = 0$$

$$\Rightarrow \frac{\vec{x} \cdot \vec{y}}{||\vec{x}|| ||\vec{y}||} = 0$$

$$\Rightarrow \vec{x} \cdot \vec{y} = 0$$

Example:

In 2D

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\vec{x} \cdot \vec{y} = 2 + (-2) = 0$$

$$|\vec{x}| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

$$|\vec{y}| = \sqrt{2^2 + (-1)^2} = \sqrt{5}$$

$$\theta = 90^\circ$$

In 3D

$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

[Open in app](#)

$$\begin{bmatrix} 5 \\ -4 \end{bmatrix}$$

$$\vec{x} \cdot \vec{y} = 2 + 18 + (-20) = 0$$

$$|\vec{x}| = \sqrt{2^2 + 3^2 + 5^2} = \sqrt{38}$$

$$|\vec{y}| = \sqrt{1^2 + 6^2 + (-4)^2} = \sqrt{53}$$

$$\theta = 90^\circ$$

In 5D

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \\ 3 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 0 \\ 1 \\ -2 \\ -3 \\ 2 \end{bmatrix} \quad \text{cos}$$

$$\vec{x} \cdot \vec{y} = 0 + 2 + (-2) + (-6) + 6 = 0$$

$$|\vec{x}| = \sqrt{1^2 + 2^2 + 1^2 + 2^2 + 3^2} = \sqrt{19}$$

$$|\vec{y}| = \sqrt{0^2 + 1^2 + (-2)^2 + (-3)^2 + 2^2} = \sqrt{18}$$

$$\cos\theta = 0$$

## Why do we care about vectors?

Let's take an example, where we have a cell phone and we want to predict whether the cell phone will hit the market or not.

[Open in app](#)

$$\begin{bmatrix} 151 & 5.8 & 1 & 1 & 3060 & 15000 & 64 \end{bmatrix}_{\mathbb{R}^{1 \times 7}}$$

This quantity has 7 dimensions.

Likewise, we can represent the data for an employee(data like tenure in months, salary, no of skills, etc.) in a vector form.



$$\begin{bmatrix} 165 & 72 & 60000 & 3 & 32 & 5 \end{bmatrix}_{\mathbb{R}^{1 \times 6}}$$

This quantity has 6 dimensions.

We can represent an input image as a vector of values where each value represents some pixel value.



$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & \dots & 1 \end{bmatrix}_{\mathbb{R}^{1 \times 784}}$$

This quantity has 784 dimensions.

We can compute the angle between two vectors and say if they are similar to each other or very far away. For example. for a new input image, we can represent it first in the form of a vector, then compute the angle between this input vector and other vectors(images in DB converted to vectors) and tell which one is most similar to the input image.

So, as obvious from the above example, we can represent each of the data points as a vector.

## Introduction to Matrices:

We can think of a matrice as a collection of many vectors.

[Open in app](#)

$$\begin{bmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

$3 \times 1 - -$

We have a vector of size  $3 \times 1$  and we have 3 such vectors stacked up, so we say we have a matrix of size  $3 \times 3$

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

$3 \times 1 - -$

$\downarrow$   
 $A \in \mathbb{R}^{3 \times 3}$

That means this matrix is denoted by values and these values are arranged in 3 rows and 3 columns.

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix}$$

[Open in app](#)

In general, we can write the dimension of a matrix as “ $m \times n$ ” where  $m$  is the no. of rows and  $n$  is the no. of columns. In the above case, both  $m$  and  $n$  equal 3.

Adding two matrices:

We can add two matrices just element-wise.

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 10 \\ 6 & 10 & 14 \\ 10 & 14 & 18 \end{bmatrix}$$

$\mathbb{R}^{m \times m}$        $\mathbb{R}^{m \times m}$        $\mathbb{R}^{m \times m}$

That means we can add up two matrices only if they have the same dimensions.

And similarly, we can subtract two matrices element-wise and in this case, also, both matrices must have the same dimensions.

### Multiplying a vector by a matrix:

Let's say we have the below matrix and we want to multiply it with a vector.

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$\mathbb{R}^{2 \times 2}$        $\mathbb{R}^{2 \times 1}$

[Open in app](#)

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} ? \\ ? \end{bmatrix}$$

Every entry in the output is the dot product of each row of the matrix with the vector.

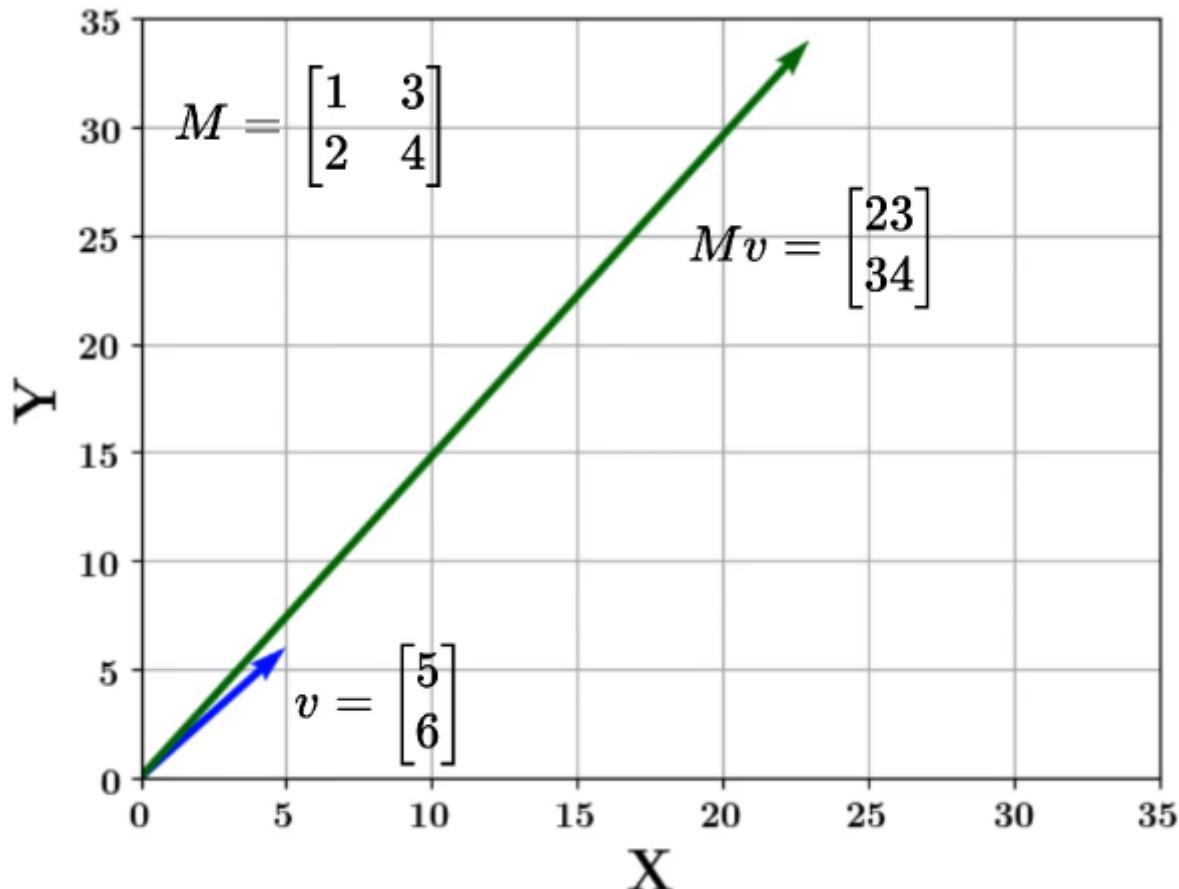
$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} (\textcircled{1} \times 5) + (\textcircled{3} \times 6) \\ ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 23 \\ (\textcircled{2} \times 5) + (\textcircled{4} \times 6) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 23 \\ 34 \end{bmatrix}$$

$2 \times 2 \quad 2 \times 1 \quad 2 \times 1$

- What happens when a matrix hits a vector?

[Open in app](#)

- The vector gets transformed into a new vector (it strays from its path).
- The vector may also get scaled (elongated or shortened) in the process.

In 3D:

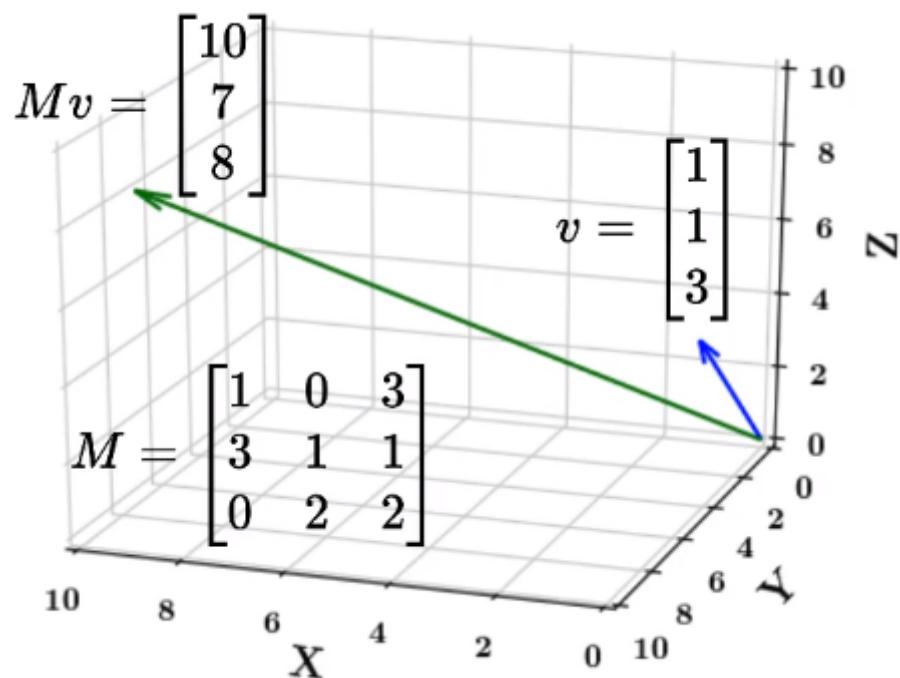
$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (\textcircled{1} \times \textcircled{1}) + (\textcircled{0} \times \textcircled{1}) + (\textcircled{3} \times \textcircled{3}) \\ ? \\ ? \end{bmatrix}$$

[Open in app](#)

$$\begin{bmatrix} 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \end{bmatrix} \begin{bmatrix} ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 7 \\ (0 \times 1) + (2 \times 1) + (2 \times 3) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 7 \\ 8 \end{bmatrix}$$




[Open in app](#)

~~matrix, it gets transformed into a new vector.~~

Now in the below case, we have a problem, because to compute the dot product, the two vectors must have the same number of elements.

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \quad \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

3x3
2x1

So, for matrix and vector multiplication, the dimensions of matrix and vector should be compatible. In other words, the number of columns in the matrix must be the same as the number of rows in the vector and the dimension of the output vector, in this case, would be the **number of rows in matrix X number of columns in the second matrix(which in this case is a vector, so second dimension is 1)**.

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} (\textcircled{1} \times \textcircled{7}) + (\textcircled{0} \times \textcircled{1}) + (\textcircled{3} \times \textcircled{?}) \\ ? \\ ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} (\textcircled{1} \times \textcircled{7}) + (\textcircled{0} \times \textcircled{1}) + (\textcircled{3} \times \textcircled{2}) \\ ? \end{bmatrix}$$

[Open in app](#)

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 7 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 13 \\ (\textcircled{3} \times \textcircled{7}) + (\textcircled{1} \times \textcircled{1}) + (\textcircled{1} \times \textcircled{2}) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 7 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 13 \\ 24 \end{bmatrix}$$

**Conclusion :** Number of columns in the matrix should be the **same** as the number of rows in the vector.



## Multiplying a matrix by another matrix

We can think of a matrix-matrix multiplication as a series of matrix-vector multiplication.

Matrix-matrix multiplication is actually the product of every row of the first matrix with every column of the second matrix. If we multiply every row of the first matrix with the first column of the second matrix, we get the first column of the output matrix.

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 5 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$$

[Open in app](#)

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 5 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & ? & ? \\ 4 & ? & ? \\ 11 & ? & ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 5 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 7 & ? \\ 4 & 5 & ? \\ 11 & 10 & ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 2 \end{bmatrix} \quad \begin{bmatrix} 5 \end{bmatrix}$$


[Open in app](#)

$$\begin{bmatrix} 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 15 \end{bmatrix}$$

Let's consider the below case where we try to multiply a '3 X 3' matrix with a '2 X 3' matrix

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 0 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \text{✖}$$



**Number of columns in the matrix is not equal to the number of rows in the vector.**

Let's look at other cases:

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 & 4 \\ 3 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \times \begin{bmatrix} 0 \end{bmatrix} = \text{✓}$$

[Open in app](#)[L](#) [-](#) [J](#)

**Number of columns in the matrix is equal to the number of rows in the vector.**

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 2 & 4 \\ 3 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 4 \\ 3 & 3 & 7 & 12 \\ 6 & 0 & 2 & 0 \end{bmatrix}$$

In general, if we have a ' $m \times n$ ' matrix and we are multiplying it with ' $n \times k$ ' matrix, we get the output as a ' $m \times k$ ' matrix.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \times \begin{bmatrix} b_{11} & \dots & b_{1k} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nk} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1k} \\ c_{21} & \dots & c_{2k} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mk} \end{bmatrix}$$

**Conclusion :** Any  $m \times n$  matrix can be multiplied with a  $n \times k$  matrix to get a  $m \times k$  output.

An alternate way of multiplying two matrices

[L](#) [-](#) [J](#)[L](#) [-](#) [J](#)[L](#) [-](#) [J](#)


[Open in app](#)

$$= b_{11} \times \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + b_{21} \times \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

This looks like the form:  $y = mx_1 + nx_2$

The output is a linear combination of the columns of the matrix and the weights in this linear combination are the elements of the vector.

Matrix-matrix multiplication can be seen as such an operation where the weights of the linear combination come from the elements of the second matrix.

## Why do we care about matrices?

We can think of the training data, test data as a matrix for example:

| Weight (g) | Screen Size(in.) | Dual SIM | Radio | Battery (mAh) | Price (INR) | Internal Memory (GB) |
|------------|------------------|----------|-------|---------------|-------------|----------------------|
| 151        | 5.8              | 1        | 1     | 3060          | 15000       | 64                   |
| 180        | 6.18             | 1        | 0     | 3500          | 32000       | 64                   |
| 160        | 5.84             | 0        | 0     | 3060          | 25000       | 64                   |
| 205        | 6.2              | 0        | 1     | 5000          | 18000       | 64                   |
| 162        | 5.9              | 0        | 1     | 3000          | 14000       | 64                   |
| 182        | 6.26             | 1        | 1     | 4000          | 12000       | 64                   |
| 138        | 4.7              | 0        | 0     | 1960          | 35000       | 64                   |
| 185        | 6.41             | 1        | 0     | 3700          | 42000       | 64                   |
| 170        | 5.5              | 0        | 0     | 3260          | 44000       | 64                   |

 $\mathbb{R}^{9 \times 7}$ 

F1      F2      F3      F4      F5      F6      Fn

$$\begin{bmatrix} \text{Smartphone icon} & 151 & 5.8 & 1 & 1 & 3060 & 15000 & 64 \\ \text{Smartphone icon} & 180 & 6.18 & 1 & 0 & 3500 & 32000 & 64 \\ \text{Smartphone icon} & 160 & 5.84 & 0 & 0 & 3060 & 25000 & 64 \end{bmatrix}$$

[Open in app](#)

|  |     |      |   |   |      |       |    |
|--|-----|------|---|---|------|-------|----|
|  | 162 | 5.9  | 0 | 1 | 3000 | 14000 | 64 |
|  | 182 | 6.26 | 1 | 1 | 4000 | 12000 | 64 |
|  | 138 | 4.7  | 0 | 0 | 1960 | 35000 | 64 |
|  | 185 | 6.41 | 1 | 0 | 3700 | 42000 | 64 |
|  | 170 | 5.5  | 0 | 0 | 3260 | 44000 | 64 |

$\mathbb{R}^{9 \times 7}$

|   | F1 | F2 | F3 | F4 | F5 | F6 | ... | Fn |
|---|----|----|----|----|----|----|-----|----|
| 0 | 0  | 1  | 1  | 0  | 0  | 0  | ... | 0  |
| 1 | 1  | 0  | 1  | 1  | 0  | 1  | ... | 1  |
| 2 | 0  | 0  | 0  | 0  | 0  | 0  | ... | 0  |
| 3 | 0  | 0  | 0  | 1  | 0  | 1  | ... | 1  |
| 4 | 0  | 1  | 0  | 1  | 1  | 1  | ... | 1  |
| 5 | 1  | 1  | 1  | 1  | 1  | 0  | ... | 0  |
| 6 | 0  | 0  | 0  | 0  | 1  | 0  | ... | 1  |
| 7 | 0  | 1  | 1  | 0  | 0  | 0  | ... | 1  |
| 8 | 1  | 0  | 0  | 0  | 1  | 0  | ... | 0  |
| 9 | 0  | 0  | 1  | 1  | 0  | 0  | ... | 0  |

$\mathbb{R}^{10 \times 784}$

And the most common operation that we would see is going to be the one given below which also involves the concept of matrix-matrix multiplication:

$$Wx + b$$

[Open in app](#) $\mathbb{R}^{m \times n}$  $\mathbb{R}^{n \times 1}$  $\mathbb{R}^{m \times 1}$ 

This article covers the content covered in the Vectors and Matrices module of the [Deep Learning course](#) and all the images are taken from the same module.

[Machine Learning](#)    [Vector](#)    [Matrix](#)    [Artificial Intelligence](#)    [Deep Learning](#)[About](#) [Write](#) [Help](#) [Legal](#)[Get the Medium app](#)